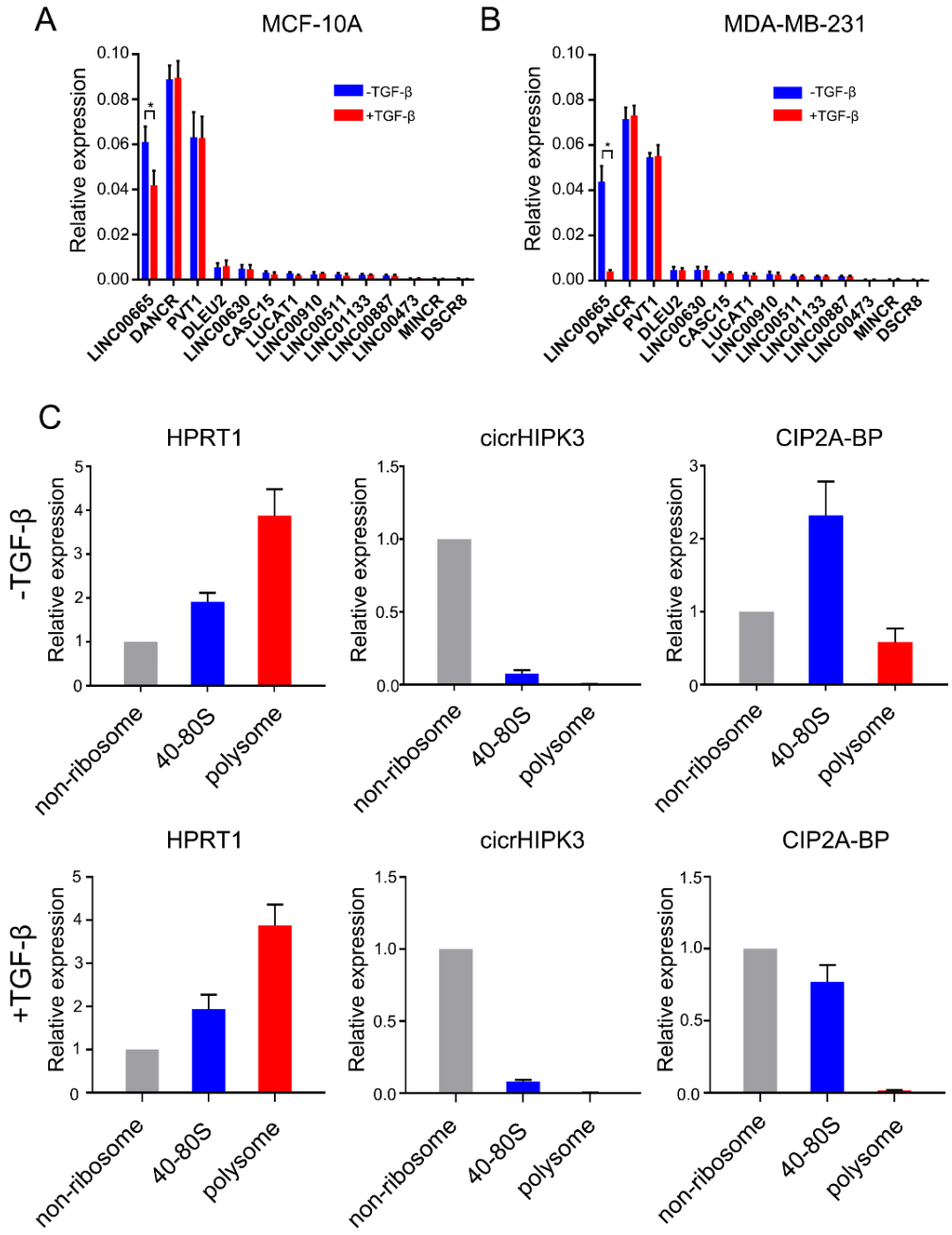
Appendix

Table of contens:



Appendix Figure S1. Screening of candidate lncRNA

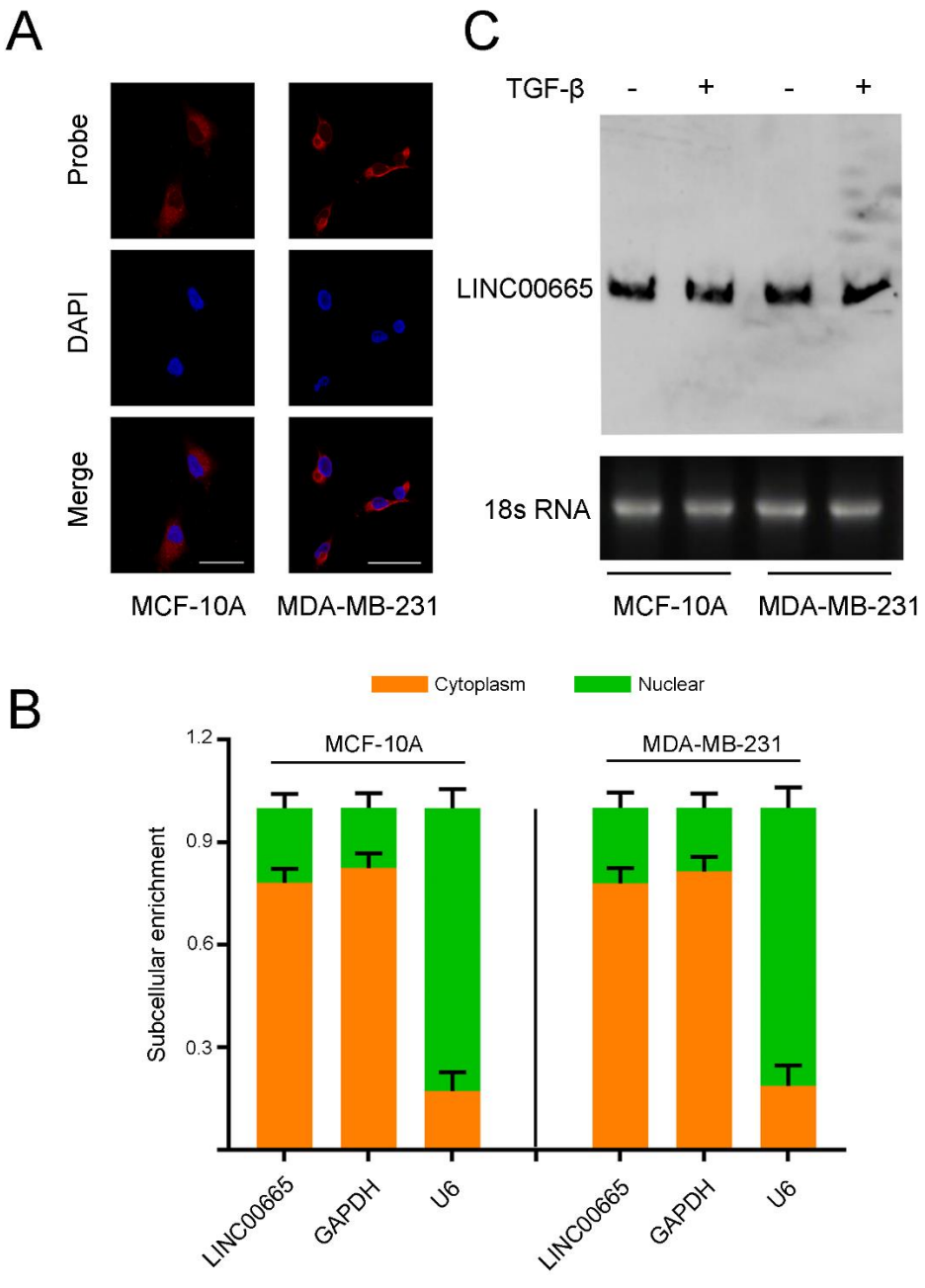
(A) Relative levels of the indicated lncRNA were determined by polysome profiling and qRT-PCR in MCF-10A either mock treated or treated with TGF-β.

(B) Relative levels of the indicated lncRNA were determined by polysome profiling and qRT-PCR in MDA-MB-231 either mock treated or treated with TGF-β.

(C) MCF-10A cell either mock treated or treated with TGF-β and whole lysates were fractionated to collect

non-ribosome, 40S-80S and polysome fractions by sucrose gradient centrifugation. HPRT1, circHIPK3 and CIP2A-BP transcripts in these fractions were quantified by qPCR.

Data information: Data are representative of 3 independent experiments (A-C). Data were assessed by paired Student's t test (A and B) and are represented as mean ± SD. *$P < 0.05$.



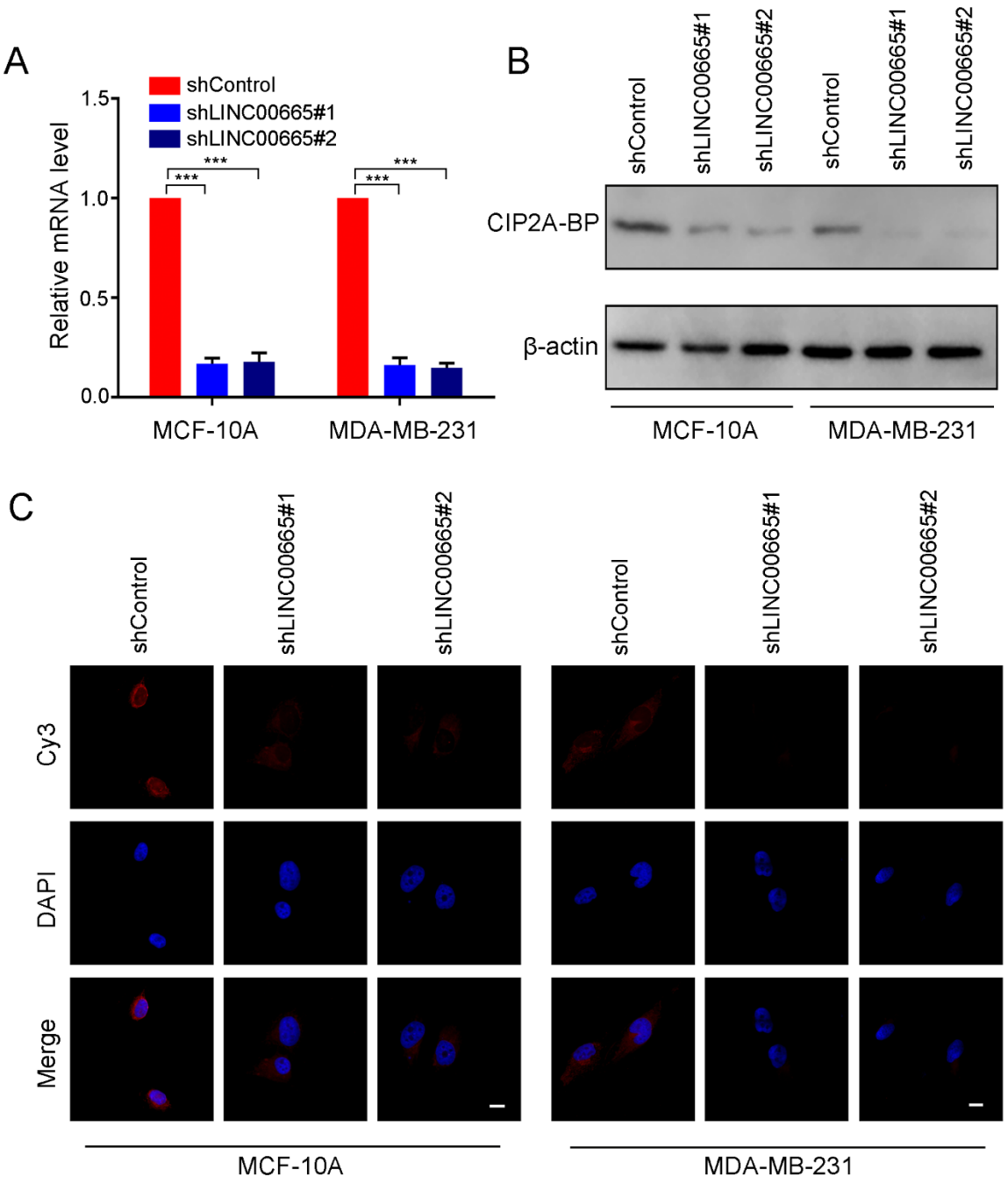
Appendix Figure S2. Biologic characterization of LINC00665

(A) RNA fluorescence in situ hybridization analysis of *LINC00665*.

(B) Relative abundance of *LINC00665* transcript in cytoplasmic and nuclear fractions from MCF-10A and MDA-MB-231 cell lines. GAPDH and U6 were used as cytoplasmic and nuclear controls respectively.

(C) Northern-blot analysis of *LINC00665* in MCF-10A and MDA-MB-231 either mock treated or treated with TGF-β.

Data information: Data are representative of 3 independent experiments (B). Scale bars: 50μm (A).



Appendix Figure S3. Anti-CIP2A-BP antibody specifically detects the micropeptide CIP2A-BP

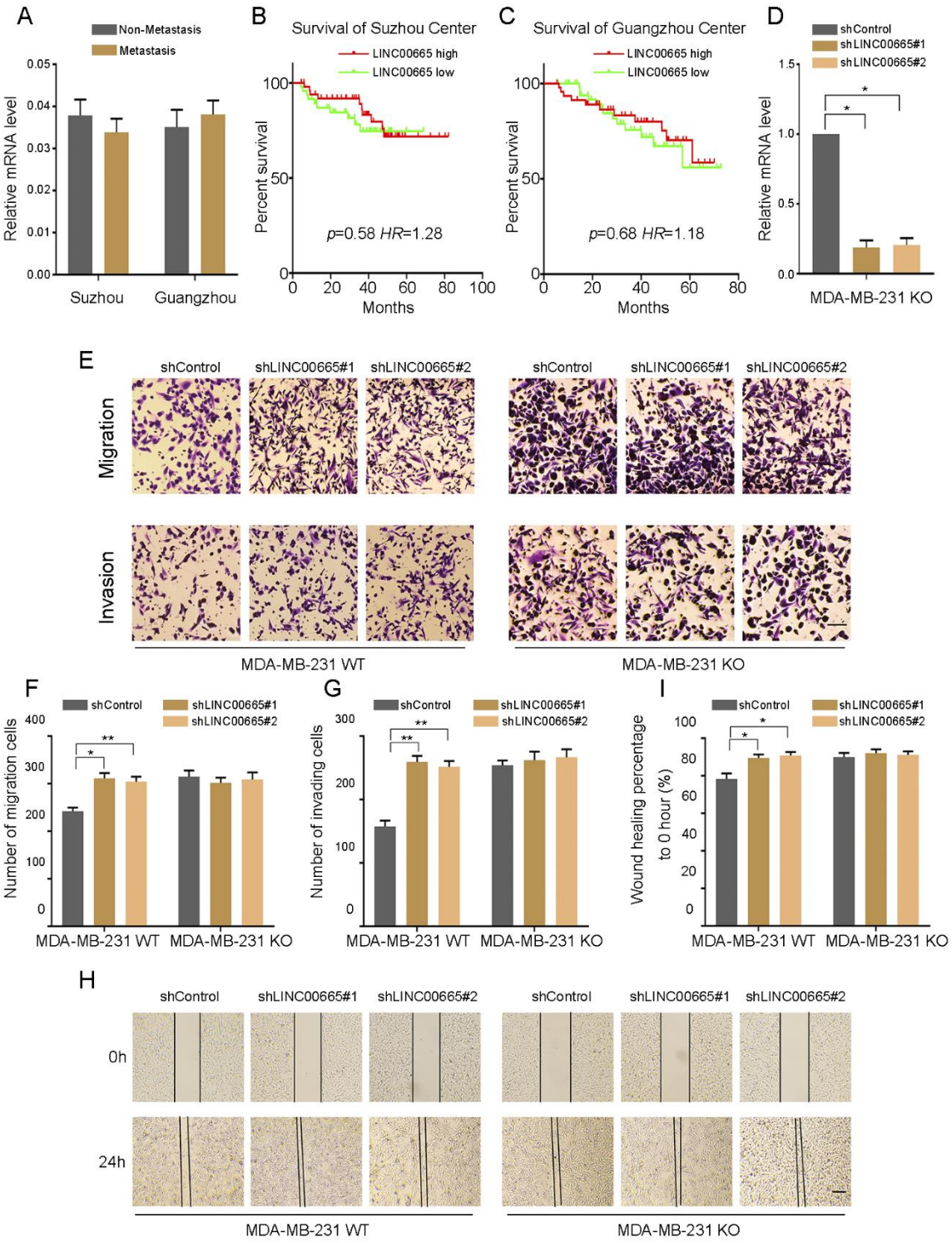
MCF-10A and MDA-MB-231 cells were transfected with *LINC00665* specific shRNA respectively.

(A) Relative *LINC00665* RNA levels were determined by qRT-PCR.

(B) Micropeptide CIP2A-BP levels were determined by western blot using anti-CIP2A-BP antibody in the indicated cells.

(C) Micropeptide CIP2A-BP was immune-stained with anti-CIP2A-BP antibody in the indicated cells.

Data information: Data are representative of 3 independent experiments (A). Data were assessed by paired Student's t test (A) and are represented as mean ± SD. ***$P < 0.01$. Scale bars, 10μm (C).



Appendix Figure S4. The *LINC00665* transcript had no significant effect on the progression of TNBC

(A) Relative levels of *LINC00665* were determined by qRT-PCR in non-metastatic and metastatic TNBC tissues in the Suzhou and Guangzhou cohorts, respectively.

(B and C) Kaplan-Meier overall survival curves for TNBC patients with high or low *LINC00665* expression in Suzhou and Guangzhou cohorts, respectively.

(D) Relative *LINC00665* RNA levels were determined by qRT-PCR.

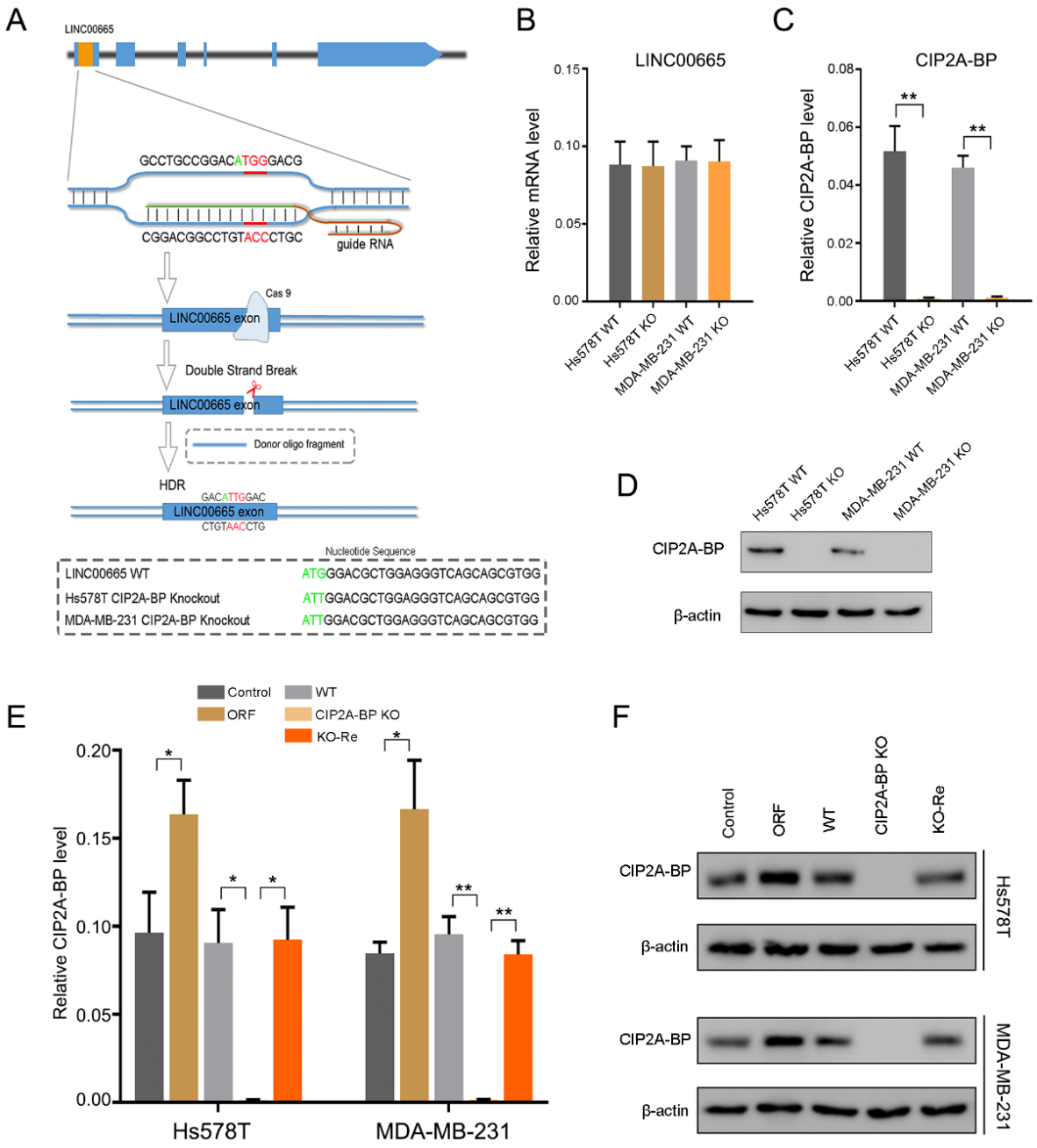
(E) Migration and invasion abilities of the indicated cells were determined using Transwell assays.

(F and G) The quantification of cells that migrated through the membrane without the Matrigel (F) and invaded through Matrigel-coated membrane (G).

(H) The indicated cells were seeded on cell culture inserts for wound-healing assay, the wound edges were photographed at the indicated time points after wounding.

(I) The quantification of the relative wound healing area.

Data information: Data are representative of 3 independent experiments (D, F, G and I). Data were assessed by paired Student's t test (D, F, G and I) and are represented as mean ± SD. Survival differences were analyzed using the log-rank test (B and C). $*P < 0.05$; $**P < 0.01$. Scale bars: 50μm (E), 200μm (H).



Appendix Figure S5. Micropeptide CIP2A-BP was stably overexpressed or knocked out

(A) Upper: diagram of generation of CIP2A-BP knock-out breast cancer cell lines using CRISP/Cas9 system. Lower: nucleotide sequence alignment of wild-type (WT) ORF 1 of human *LINC00665* and representative mutated *LINC00665* forms.

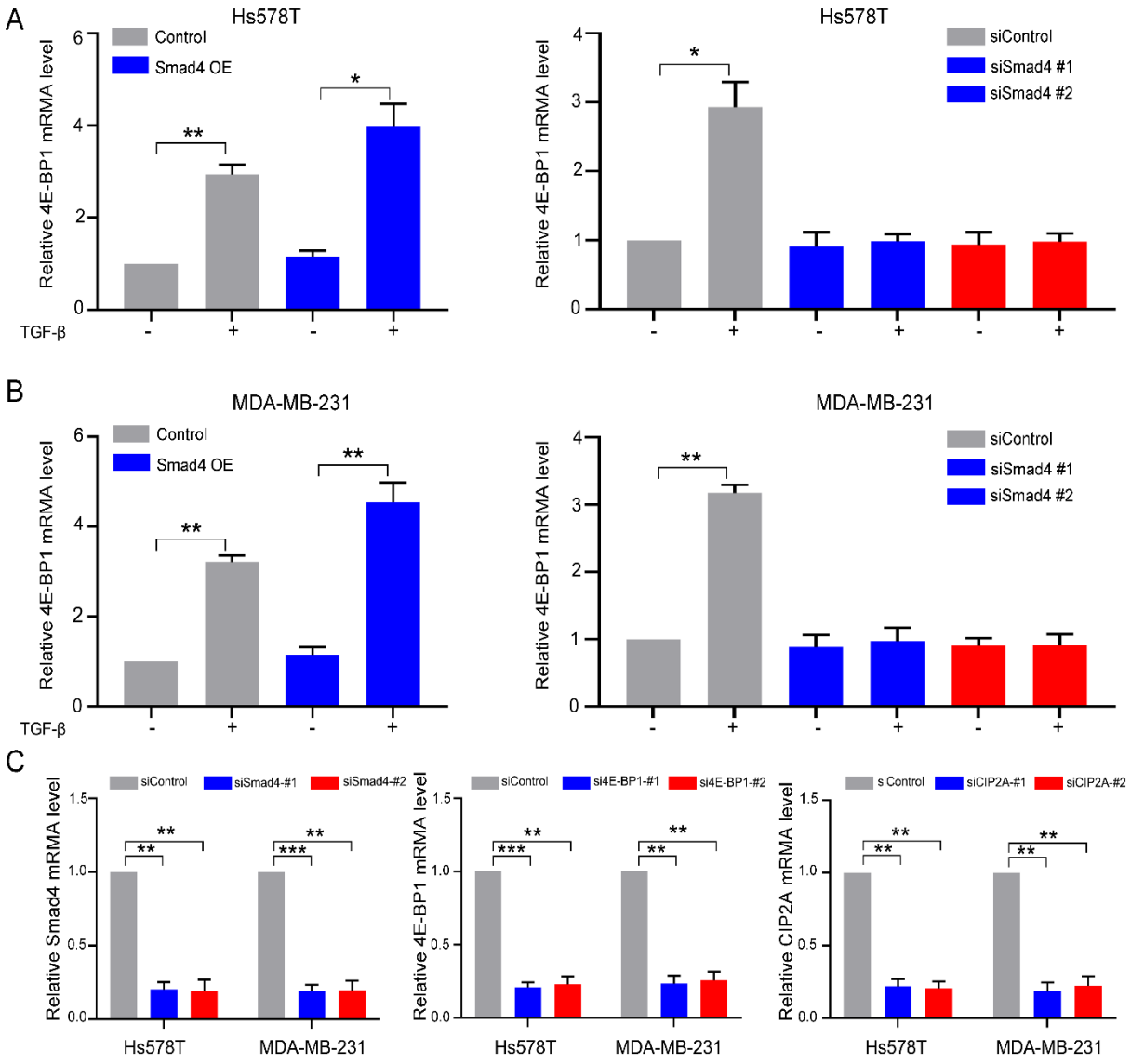
(B) Relative *LINC00665* RNA levels were determined by qRT-PCR in the indicated cells.

(C and E) Relative levels of micropeptide CIP2A-BP were determined by polysome profiling and qRT-PCR

in the indicated cells.

(D and F) Micropeptide CIP2A-BP levels were determined by western blot using anti-CIP2A-BP antibody in the indicated cells.

Data information: Data are representative of 3 independent experiments (B-E). Data were assessed by paired Student's t test (B-E) and are represented as mean ± SD. **P* < 0.05; ***P* < 0.01.

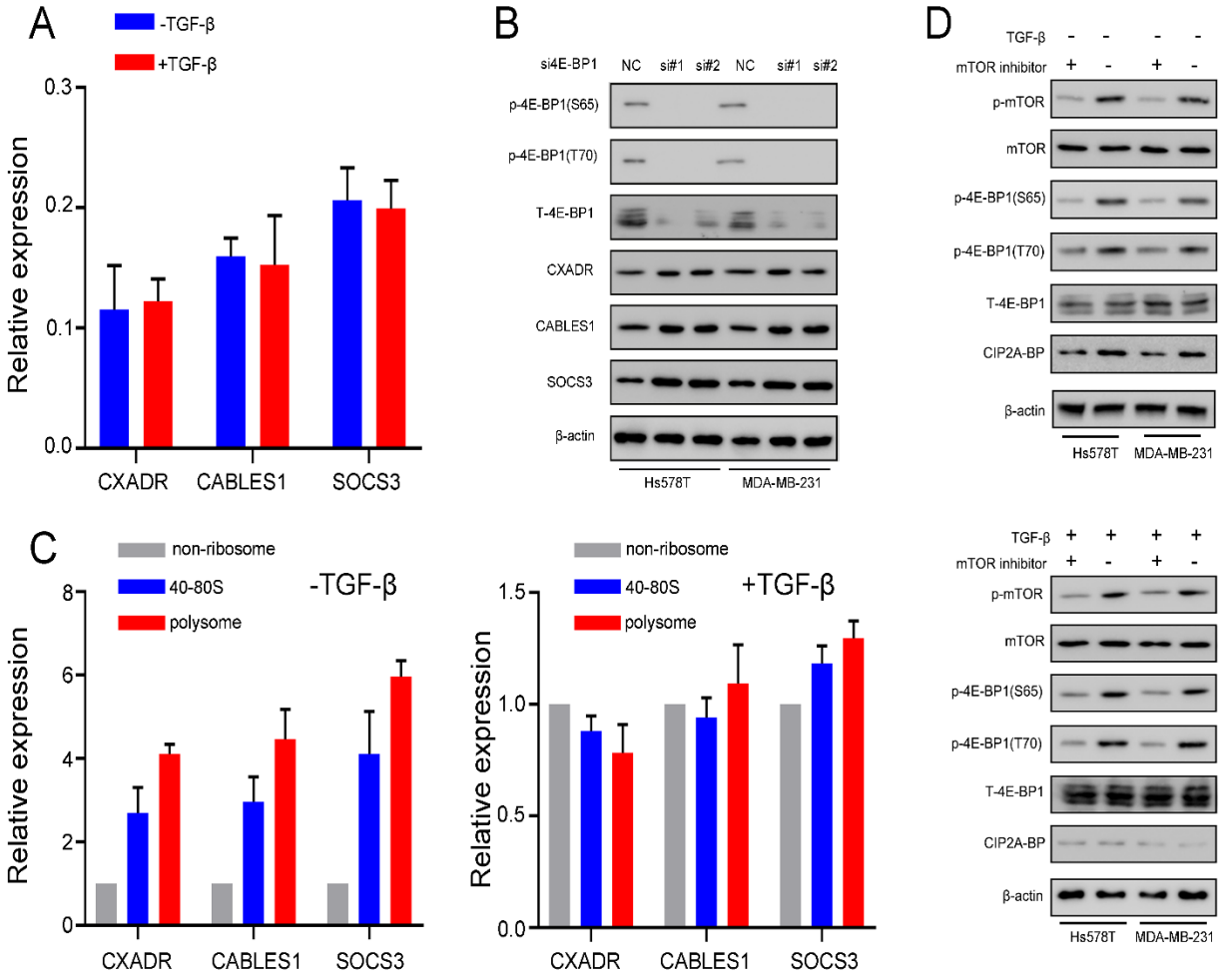


Appendix Figure S6. TGF-β/Smad4 signaling pathway increased endogenous 4E-BP1 mRNA level

(A and B) Endogenous *4E-BP1* mRNA is elevated by TGF-β and Smad4 in Hs578T(A) and MDA-MB-231(B) cells.

(C) The mRNA levels of *Smad4*, *4E-BP1* and *CIP2A* in Hs578T and MDA-MB-231 cells that transfected with specific siRNA are significantly lower than that in cells transfected with empty vector.

Data information: Data are representative of 3 independent experiments (A, B and C). Data were assessed by paired Student's t test and are represented as mean ± SD. **P* < 0.05; ***P* < 0.01; ****P* < 0.001.



Appendix Figure S7. 4E-BP1 affect the translation of multiple mRNAs

By analyzing the original data of this study (GSE59817), 429 genes were identified, whose translation but not transcription was decreased significantly by TGF-β treatment (logFC<-1, FDR<0.05).

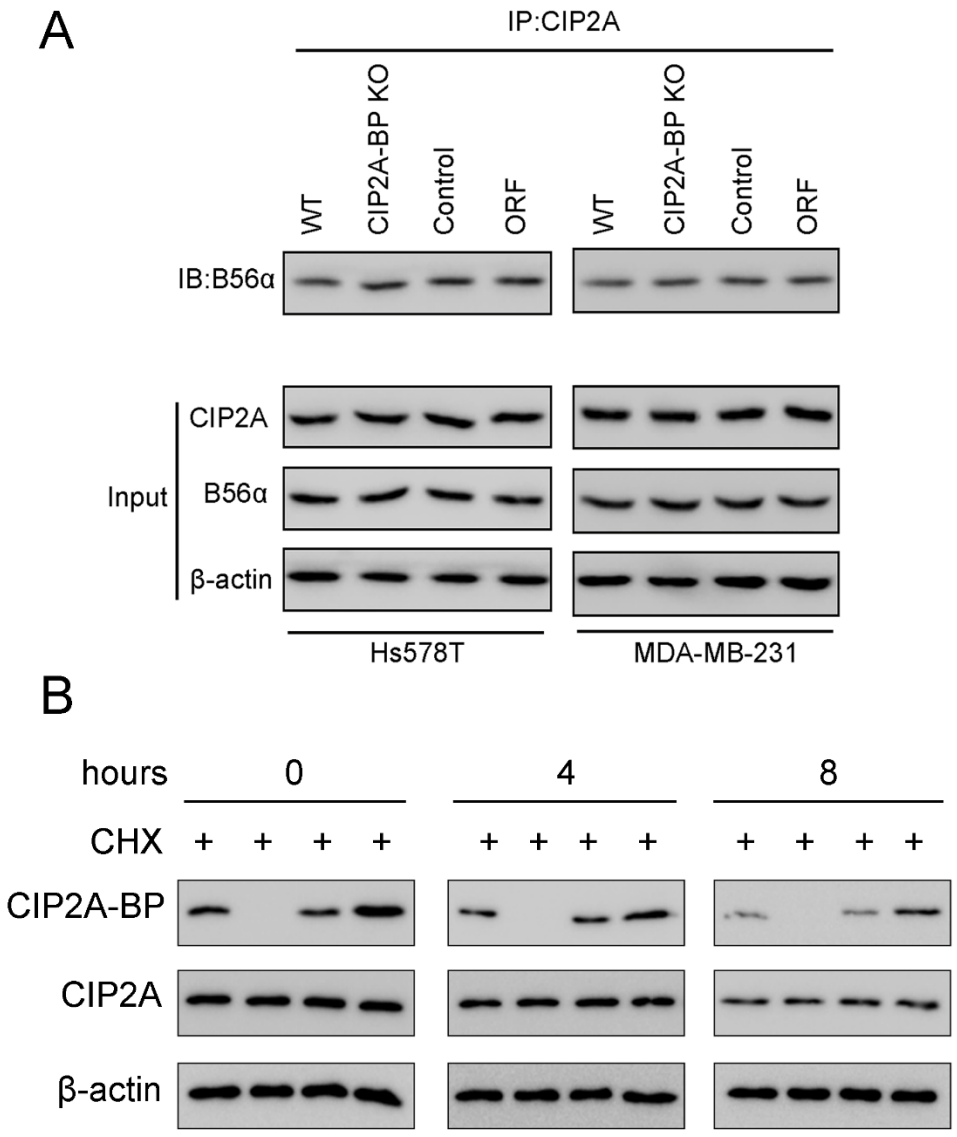
(A) Relative levels of CXADR, CABLES1 and SOCS3 mRNA were determined by qRT-PCR in MCF-10A either mock treated or treated with TGF-β.

(B) Downregulation of 4E-BP1 by siRNA treatment significantly upregulated protein levels of CXADR, CABLES1 and SOCS3 determined by western blot analysis.

(C) TGF-β treatment of MCF-10A cell significantly reduced the percentage of CXADR, CABLES1 and SOCS3 mRNA in the 40-80S and polysome fractions.

(D) TNBC cells were pretreated with specific antagonist against mTOR (PP242, 1μM) for 1h and then cultured with or without TGF-β. The indicated proteins were determined by immunoblotting analysis.

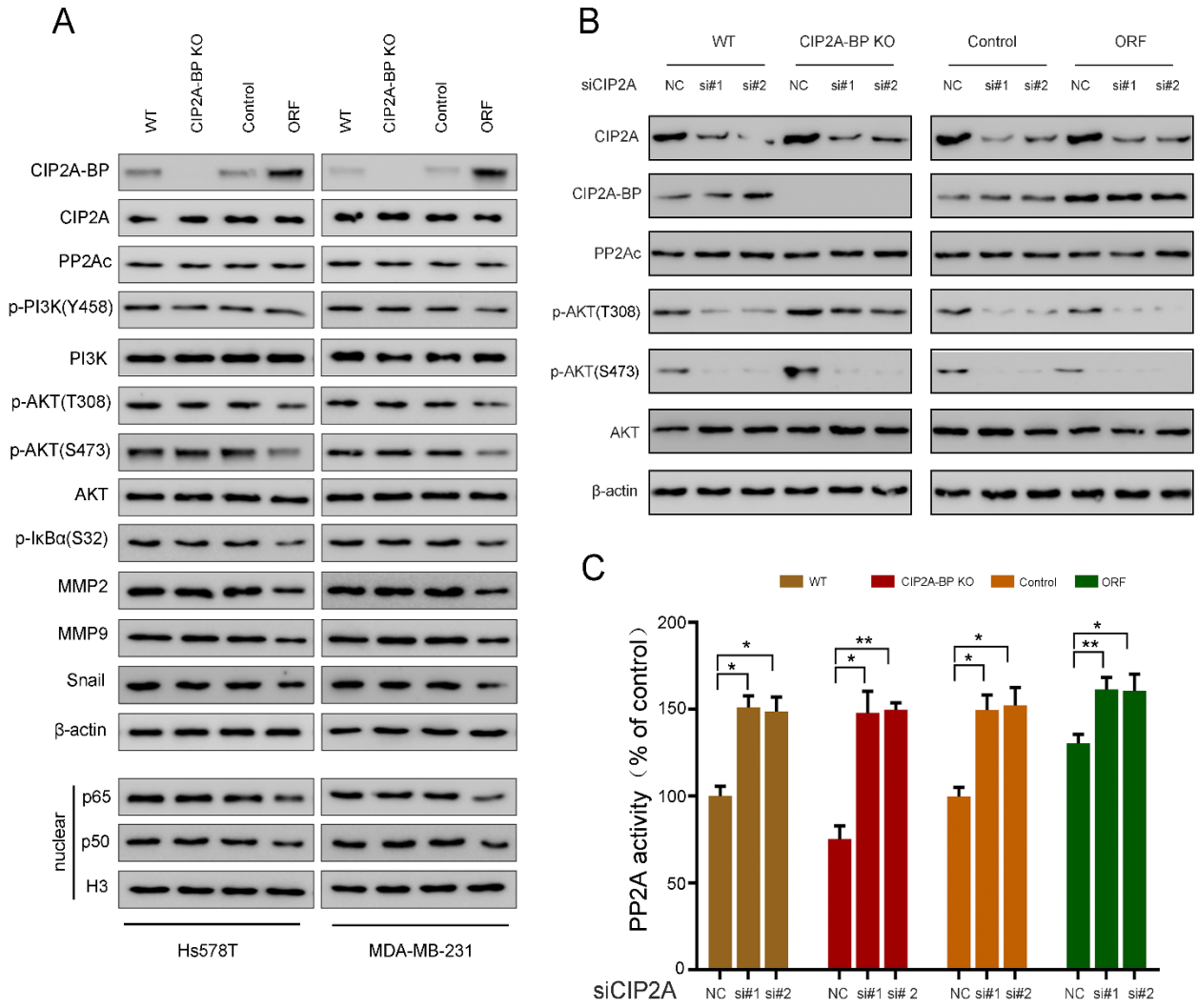
Data information: Data are representative of 3 independent experiments (A and C). Data were assessed by paired Student's t test (A) and are represented as mean ± SD.



Appendix Figure S8. Micropeptide CIP2A-BP does not affect the stability of CIP2A and its binding to the B56α subunit of PP2A

(A) Co-immunoprecipitation assays showed that the interaction between CIP2A and B56α in TNBC cells was unaffected by CIP2A-BP expression.

(B) After cells were treated with translation inhibitor (CHX, 100μg/ml) for the indicated periods, the stability of CIP2A protein in whole-cell lysates was assessed by immunoblotting analysis.



Appendix Figure S9. Related to Figure 6

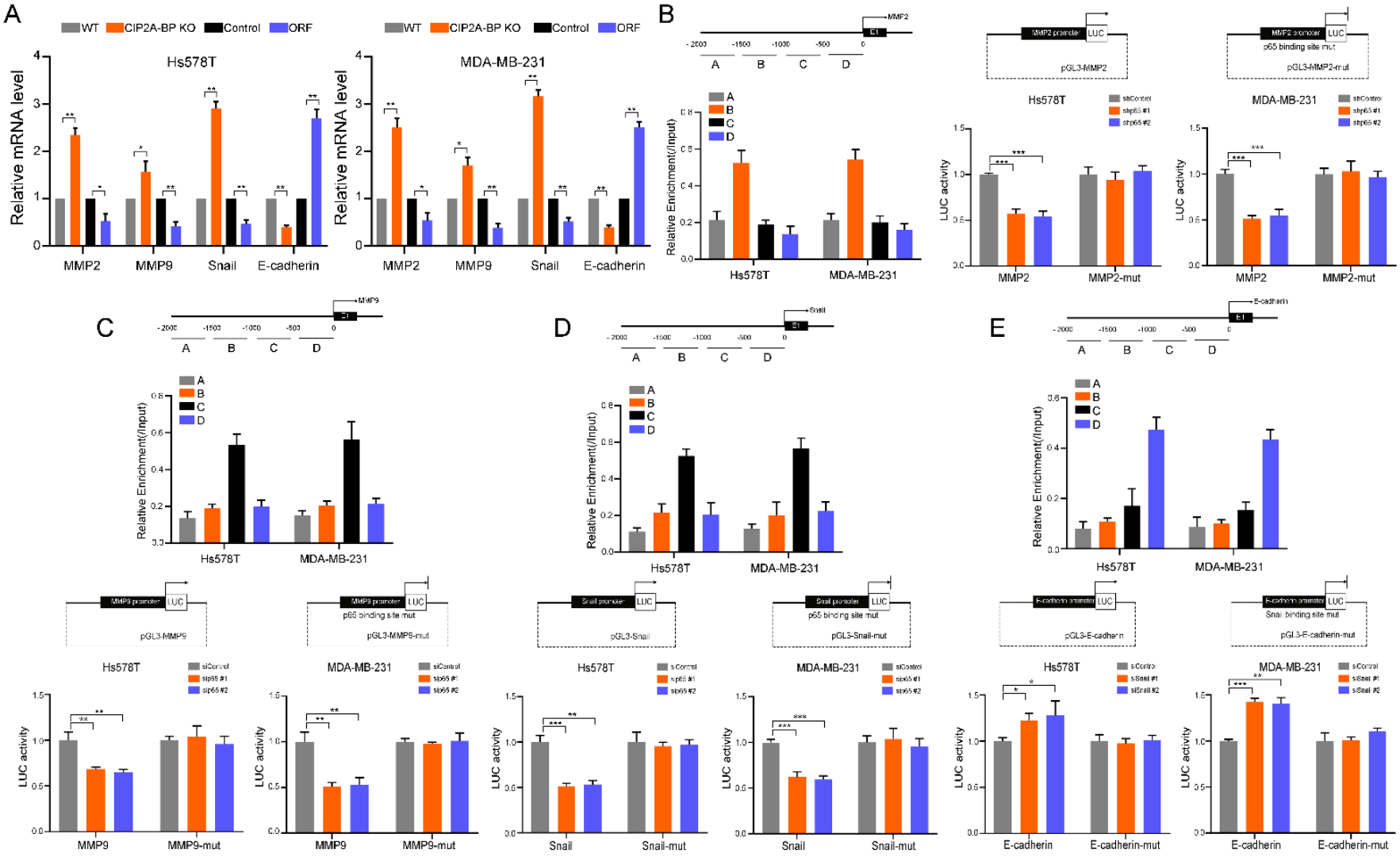
(A) The expression levels of the indicated proteins were detected by immunoblotting analysis in CIP2A-BP KO, *LINC00665* ORF overexpressed (OE) and respective controls of TNBC cells treated with TGF-β.

(B) Hs578T cells were transfected with anti-CIP2A siRNAs. Whole-cell lysates of these cells were subjected

to immunoblot analysis with CIP2A, CIP2A-BP, PP2Ac, p-AKT, AKT and β-actin antibodies.

(C) CIP2A-BP KO, *LINC00665* ORF overexpressed (OE) and respective control Hs578T cells were transfected with anti-CIP2A siRNA. PP2A activity assay was performed to evaluate PP2A activity in the indicated cells.

Data information: Data are representative of 3 independent experiments (C). Data were assessed by paired Student's t test and are represented as mean ± SD. $*P < 0.05$; $**P < 0.01$.



Appendix Figure S10. Micropeptide CIP2A-BP affects the transcription of *MMP2, MMP9, Snail* and *E-cadherin* genes through PI3K/AKT/ NFκB

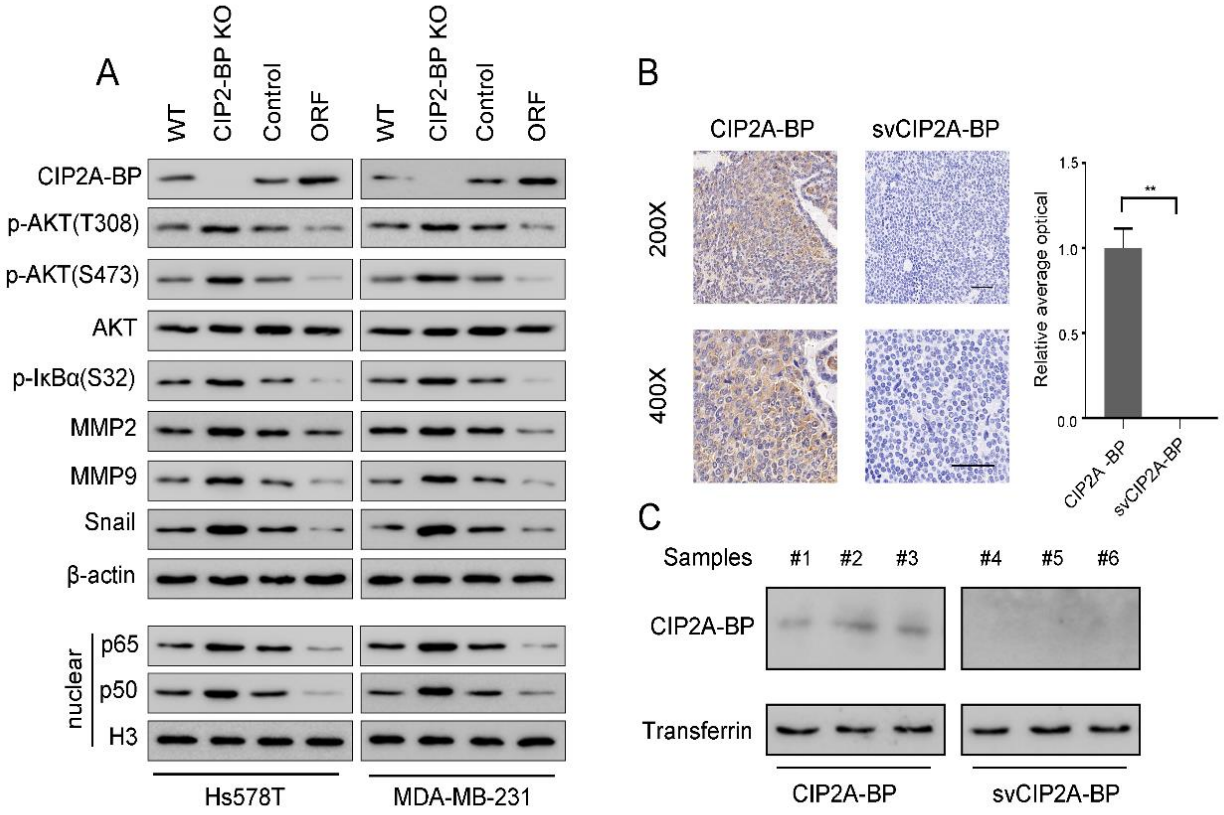
signaling pathways

(A) Relative *MMP2, MMP9, Snail* and *E-cadherin* RNA levels were determined by qRT-PCR.

(B-D) Chromatin immunoprecipitation showing p65 occupancy at the MMP2 (B), MMP9 (C) and Snail (D) promoter region, and MMP2, MMP9 and Snail promoter activity levels were measured using the luciferase assay in TNBC cells.

(E) Chromatin immunoprecipitation showing Snail occupancy at the E-cadherin promoter region, and E-cadherin promoter activity levels were measured using the luciferase assay in TNBC cells.

Data information: Data are representative of 3 (A-E) or 4 (B-G) independent experiments. Data were assessed by paired Student's t test (A-E) and are represented as mean ± SD. **P* < 0.05; ***P* < 0.01; ****P* < 0.001.



Appendix Figure S11. Related to Figure 6 and Figure 7

(A) Immunoblotting analysis of expression of the indicated proteins in TNBC cells untreated with specific

antagonist against AKT (MK-2206).

(B) MMTV-PyMT mice were injected with CIP2A-BP or svCIP2A-BP via mammary fat pad. Immunohistochemistry (IHC) staining of CIP2A-BP of mammary tumors.

(C) The levels of CIP2A-BP were detected by immunoblotting analysis in serum samples of mice injected with CIP2A-BP or svCIP2A-BP via tail vein.

Data information: Data are representative of 3 independent experiments (B). Data were assessed by paired Student's t test (B) and are represented as mean ± SD. $^{**}P < 0.01$. Scale bars: 50μm (B).

Appendix Table S1. Primary clinical and histological characteristics of 217 study patients

Characteristics	Suzhou population	Guangzhou population	Overall
	Patients, n (%)	Patients, n (%)	Patients, n (%)
Age (years)			
≤40	23 (20.5)	34 (32.4)	57 (26.3)
41–60	77 (68.8)	67 (63.8)	144 (66.3)
≥60	12 (10.7)	4 (3.8)	16 (7.4)
Age at menarche (years)			
≤14	57 (50.9)	52 (49.5)	109 (50.2)
>14	55 (49.1)	53 (50.5)	108 (49.8)
Menopausal history			
Pre-menopausal	77 (68.8)	49 (46.7)	126 (58.1)
Menopausal	35 (31.2)	56 (53.3)	92 (41.9)
Body mass index			
≤20	36 (32.1)	24 (22.9)	60 (27.6)
20 < BMI < 28	71 (63.4)	73 (69.5)	144 (66.4)
≥28	5 (4.5)	8 (7.6)	13 (5.0)
Family history			
Positive	4 (3.6)	11 (10.5)	15 (6.9)
Negative	108 (96.4)	94 (89.5)	202 (93.1)
Pathological type			
Invasive ductal carcinoma	105 (93.8)	97 (92.4)	202 (93.1)
Others carcinoma	7 (6.2)	8 (7.6)	15 (6.9)
Smoking			
Positive	40 (35.7)	27 (25.7)	67 (30.9)
Negative	72 (64.3)	78 (74.3)	150 (69.1)
Drinking			
Positive	78 (69.6)	63 (60.0)	141 (65.0)
Negative	34 (30.4)	42 (40.0)	76 (35.0)
Stage			
I	25 (22.3)	24 (22.9)	49 (22.6)
II	42 (37.5)	37 (35.2)	79 (36.4)
III	31 (27.7)	29 (27.6)	60 (27.6)
IV	14 (12.5)	15 (14.3)	29 (13.4)
Lymph node invasion			
Positive	56 (50.0)	49 (46.7)	105 (48.4)
Negative	56 (50.0)	56 (53.3)	112 (51.6)

Appendix Table S2. Virus Strains and Recombinant DNA Used in This Study

Virus Strains	Source	Identifier
Lentivirus pLVX-IRES-neo	Genepharma	N/A

Lentivirus pLVX-IRES-neo-LINC00665 ORF-His	Genepharma	N/A
pcDNA3.1(+)	Invitrogen	V790-20
ORF-His	This paper	N/A
ORFmut-His	This paper	N/A
ORF2-His	This paper	N/A
ORF3-His	This paper	N/A
ORF4-His	This paper	N/A
pEGFP-N1 (GFPwt)	Clontech	6085-1
GFPmut	This paper	N/A
ORF-GFPmut	This paper	N/A
ORFmut-GFPmut	This paper	N/A

Appendix Table S3. Antibodies

Antibodies	**Source**	**Identifier**
Mouse Monoclonal anti-His	Santa Cruz	sc-53073
Mouse Monoclonal anti-GFP	Santa Cruz	sc-9996
Rabbit Polyclonal anti-CIP2A-BP	This paper	N/A
Rabbit polyclonal anti-β-actin	Abcam	Ab8227
Rabbit polyclonal anti-E-Cadherin	BosterBio	PB9561
Rabbit polyclonal anti-N-Cadherin	BosterBio	PA1328
Rabbit Monoclonal anti-Phospho-Smad2 (Ser465/467)	Cell Signaling Technology	3108
Rabbit Monoclonal anti-Smad2	Cell Signaling Technology	5339
Rabbit Monoclonal anti-Phospho-Smad3 (Ser423/425)	Cell Signaling Technology	9520
Rabbit Monoclonal anti-Smad3	Cell Signaling Technology	9523
Rabbit Monoclonal anti-Smad4	Cell Signaling Technology	38454
Rabbit Polyclonal anti- Phospho-4EBP1 (Ser65)	Cell Signaling Technology	9451
Rabbit Polyclonal anti- Phospho-4EBP1 (Thr70)	Cell Signaling Technology	9455
Rabbit Polyclonal anti-eIF4EBP1	Cell Signaling Technology	9452
Rabbit Polyclonal anti-eIF4E	Abcam	Ab1126
Rabbit Monoclonal anti-Filamin A	BosterBio	BM4039
Rabbit Polyclonal anti-CIP2A	Abcam	Ab99518
Rabbit Polyclonal anti-ANXA2	BosterBio	BA0641
Rabbit Polyclonal anti-MYL6	Abcam	Ab174169
Rabbit Polyclonal anti-CLTC	Bioss	bs-2932R
Mouse Monoclonal anti-PFKP	Abcam	Ab119796
Rabbit Polyclonal anti-Myc tag	Abcam	Ab9106
Mouse Monoclonal anti-PP2A-B56α	Santa Cruz	sc-136045
Mouse Monoclonal anti-PP2A-B56γ	Santa Cruz	sc-374380
Mouse Monoclonal anti-PP2Ac	BD Biosciences	610555

Rabbit Polyclonal anti- Phospho-PI3K(Tyr458)	Invitrogen Antibodies	PA5-17387
Rabbit Monoclonal anti- PI3K	Cell Signaling Technology	4257
Rabbit Monoclonal anti-Phospho-Akt (Thr308)	Cell Signaling Technology	4056
Mouse Monoclonal anti-Phospho-Akt (Ser473)	Cell Signaling Technology	4051
Mouse Monoclonal anti-AKT	Cell Signaling Technology	2920
Rabbit Monoclonal anti-Phospho-IκBα (Ser32)	Abcam	Ab92700
Rabbit Polyclonal anti-p65	Abcam	Ab16502
Rabbit Monoclonal anti-p50	Abcam	Ab32360
Rabbit Polyclonal anti-MMP2	Abcam	Ab97779
Rabbit Polyclonal anti-MMP9	Abcam	Ab73734
Mouse Monoclonal anti-Snail	Cell Signaling Technology	3895
Rabbit Monoclonal anti-H3	Cell Signaling Technology	4499
Rabbit Polyclonal anti-Phospho-mTOR (Ser2448)	Cell Signaling Technology	2971
Rabbit Polyclonal anti-mTOR	Cell Signaling Technology	2972
Rabbit Polyclonal anti-CXADR	Abcam	Ab100811
Rabbit Polyclonal anti-CABLES1	Abcam	Ab134258
Rabbit Polyclonal anti-SOCS3	Abcam	Ab16030
Rabbit Polyclonal anti-Transferrin	Abbkine	ABP52968
Goat anti-Mouse IgG (H+L) HRP	Santa Cruz	sc-2005
Goat anti-Rabbit IgG (H+L) HRP	Santa Cruz	sc-2004

Appendix Table S4. Primers Used for qPCR

	Forward primers sequence	**Reverse primers sequence**
LINC00665	AGGTACCCACCAGATGTGTCT	ACCACAAGAAAGGTGGATGGA
DANCR	GTTCTTAGCGCAGGTTGAC	AACTGAAGGGATAGTTGGCT
PVT1	CTGACTTCGCTGATTAAGTGG	AGACTGGCTCATCAGATGG
DLEU2	CTTAAGTCTGTTACTTGGATTACGG	ATTATCCACTTCCTGGATACTCTC
LINC00630	AGATGGGAATCACATTGCTC	ACAGATTACCTGGCAAAGAG
CASC15	CTACAGATGTGTTAACACCCAG	ATTGCCATGATTCACACCTG
LUCAT1	CAATGGTATTTCTGACTTGGCT	AGCGAAACTCTGTAGCTCAG
LINC00910	TAGAACTGGTATCTCTTCAACTACC	CCTGTCTTCTGGAATGGCT
LINC00511	CTTCTCCTGTAACGTGTGG	CTGATTCCACCACGTTCTC
LINC01133	AGGAGCCATTAACAAAGCT	GAATCTTCAGTTGCAGGGA
LINC00887	GTCTCCATTTCACAGATGAGAG	GTTTAAGAGGAGGCTGCTG
LINC00473	TAGAACATCACCTTGGTGC	TCCCATATTTCCTGTAAGCTC
MINCR	GGCAAGAGCACTTCCTTCCG	TCTTCTGCGAAGTTGGGAACC
DSCR8	CACTCGTTTCTGTAAACATGG	CAAGATGAAGCTGGGATTAGG
LINC00665 ORF-His	GTGGAGTCCTGGCCTTTTG	GGTGGACGGATGAGAAACGG
HPRT1	CTTTGCTGACCTGCTGGATT	TCCCCTGTTGACTGGTCATT
circHIPK3	TCGGCCAGTCATGTATCAAA	ACCAAGACTTGTGAGGCCAT
β-actin	CATGTACGTTGCTATCCAGGC	CTCCTTAATGTCACGCACGAT

Appendix Table S5. Primers Used for Northern

	Forward primers sequence	Reverse primers sequence
LINC00665	GGTGCAAAGTGGGAAGTGTG	TAATACGACTCACTATAGGG TCCGGTGGACGGATGAGAAA

Appendix Table S6. Primers Used for ChIP

	Forward primers sequence	Reverse primers sequence
4E-BP1-A	CTCGGTCTCCAAATGTGCTG	GGAGGCTGAGGTGGAGAATT
4E-BP1-B	GTGTTGGGATTACAGGCGTG	GCTACAGGTGCACATTGAGC
4E-BP1-C	AGGGTCTCTGTCACAAAGGG	GAAGAGGGTGAAGAGGGGAG
4E-BP1-D	TTGGACTCAATTCAGGGGCT	TTAGAGAGGGTGTGGGAGGA
4E-BP1-E	CCCAACCCAGCCAAGGATTA	GTCTGCTCACGGGCTCTT
MMP2-A	AATCGTGACCTCCAATGCCC	GGAAGGGGGCAGATAGGACT
MMP2-B	AGTCAGGCGTTCCCAACAG	TTGAATCCTTTCCTGCGCCA
MMP2-C	TAAGGAAGCAACCTGGGACC	GGAACGGGAAGGAATGGTCA
MMP2-D	GCACTGAGGGTGGACGTAGAG	GATGCAGCGGAAACAAGGGAG
MMP9-A	ATCTCCATCTCACAGTCTCATTT	ATTTAAAGGGCCTACTATGTGCC
MMP9-B	TATAGACCCTGCCCGATGCC	GCCAAATCTCCAGCCCCAAT
MMP9-C	TCTTGGGTCTTGGCCTTAGT	GTCTTCCGCAGGCTGAATCT
MMP9-D	ACTTTCCCTTGGCTGACCAC	AAACTGCAGAGCTTGTGGGA
Snail-A	AGAATGAAAGGAAGCCAGCGT	CAAAATCGTCCTCCCCTGGTT
Snail-B	CGGCACCAAGTGACTAAACAG	CCTCACAGGTCTCACCGTTC
Snail-C	CTCGGGCCTTTTCCCTTGAT	TTGACGAGGGAAACGCACAT
Snail-D	CGTCGGAAGGTCAGGTGTC	TAGCGCAGAAGAACCACTCG
E-cadherin-A	GCTGGTCAGTGTCAAATGCT	AGAATGGCATGAACCTGGGA
E-cadherin -B	GGAGAAACTGAGGCTTTGGGA	CTTCATGGGTTAGTGAGTCAGC
E-cadherin -C	TGGCTCACACCTGAAATCCT	GCCTCTCTAGTAGCTGGGAG
E-cadherin -D	CTCAGCCAAGTGTAAAAGCCC	TGCACGGTTCTGATTCCACT

Appendix Table S7. Sequences of siRNAs and shRNAs Used in This Study

	No.	Sequence (5'-3')
LINC00665	shLINC00665-#1 shLINC00665-#2	5’-GTGGAATGTCAGTGACTTAGC-3’ 5’-GCCTCCCTAGAGACCATTTCT-3’
Smad4	siSmad4-#1 siSmad4-#2	5’-GCCAUAGUGAAGGACUGUU-3’ 5’-GCUCCUAGACGAAGUACUU-3’
4E-BP1	si4E-BP1-#1 si4E-BP1-#2	5’-CGAACCCUUCCUUCCGAAUUU-3’ 5’-GAUCAUCUAUGACCGGAAA-3’
CIP2A	siCIP2A-#1 siCIP2A-#2	5’-GGACGUGGUUUACUGCGAU-3’ 5’-GCGUAAAUACCUCGCUAAA-3’

P65	sip65-#1 sip65-#2	5’-GCUGCAGUUUGAUGAUGAAGA-3’ 5’-UGGAUUCAUUACAGCUUAAUC-3’
Snail	siSnail-#1 siSnail-#2	5’-GCAGUUAAUUUAUAUAUUAAA-3’ 5’-GUUUAUUGAUAUUCAAUAAAG-3’